



<http://www.diva-portal.org>

This is the published version of a paper published in *Education Inquiry*.

Citation for the original published paper (version of record):

Hanberger, A., Lindgren, L., Lundström, U. (2016)

Navigating the evaluation web: evaluation in Swedish local school governance.

*Education Inquiry*, 7(3): 259-281

<http://dx.doi.org/10.3402/edui.v7.29913>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-121157>

# Navigating the Evaluation Web: Evaluation in Swedish Local School Governance

Anders Hanberger\*<sup>†</sup>, Lena Lindgren\*\* & Ulf Lundström\*\*\*

## Abstract

This paper explores the use, functions and constitutive effects of evaluation systems in local school governance, and identifies how contextual factors affect various uses of evaluation in this context. This case study of three Swedish municipalities demonstrates that local evaluation systems are set up to effectively sustain local school governance and ensure compliance with the Education Act and other state demands. Local decision makers have learned to navigate the web of evaluations and developed response strategies to manage external evaluations and to take into account what can be useful and what cannot be overlooked in order to avoid sanctions. The study shows that in contexts with high issue polarisation, such as schooling, the use of evaluation differs between the political majority and opposition, and relates to how schools perform in national comparisons and school inspections. Responses to external evaluations follow the same pattern. Some key performance indicators from the National Agency of Education and the School Inspectorate affect local school governance in that they define what is important in education, and reinforce the norm that benchmarking is natural and worthwhile, indicating constitutive effects of national evaluation systems.

*Keywords:* evaluation systems, local school governance, context, navigation, use, constitutive effects

## 1. Introduction

Schools are increasingly being monitored by school inspections and benchmarked by national and transnational evaluation systems<sup>1</sup> (Levin 2010; Lingard and Sellar 2013; Merki 2011; Ravitch 2010). This paper demonstrates that, to manage this growing accountability pressure, local politicians and school administrators have developed response strategies and set up their own evaluation systems to maintain and strengthen local school governance<sup>2</sup>. Like many countries, Sweden is ever more emphasising evaluation at all levels of the education system (OECD 2011). In the last decade, more than 20 educational reforms have been implemented, including a more distinct system of student assessment, national tests starting at earlier ages and in a greater number of subjects, and the establishment of a Swedish Schools Inspectorate

---

<sup>†</sup>Correspondence to: Anders Hanberger, Department of Applied Educational Science, Umeå University, Umeå, Sweden. Email: anders.hanberger@umu.se

\*Department of Applied Educational Science, Umeå University, Sweden. Email: anders.hanberger@umu.se

\*\*School of Public Administration, Gothenburg University, Sweden. Email: lena.lindgren@spa.gu.se

\*\*\*Department of Applied Educational Science, Umeå University, Sweden. Email: ulf.p.lundstrom@umu.se

©Authors. ISSN 2000-4508, pp. 259–281

(SSI) with substantially enhanced sanctioning tools. This paper explores how Swedish local decision makers navigate the education evaluation web, focusing on evaluation in *the local government model of governance* (Hanberger, in this issue). As little previous research treats this matter, this paper is explorative in investigating the role of evaluation in school governance in three municipalities.

Swedish schools are monitored, benchmarked and held accountable through various evaluation systems (Lindgren, Hanberger and Lundström, in this issue). Most evaluation systems are ultimately intended to enhance performance and ensure quality in education, although their methods and mechanisms for achieving this differ. The SSI monitors compliance with laws and regulations and imposes sanctions to enhance compliance, assuming that this will improve education (de Wolf and Janssens 2007; Ehren, Altrichter, McNamara and O'Hara 2013; Leeuw 2003; Whitby 2010). In contrast, benchmarking systems develop new or use existing performance measures, assuming that comparing and ranking schools will lead to school improvement (Andersen, Dahler-Larsen and Strömbæk Pedersen 2009; Elstad 2009; Lingard and Sellar 2013; Mintrop and Trujillo 2007). Local education committees are supposed to use school inspection reports and evaluation results when deciding on feasible action to develop schools and enhance quality in education. Low-ranked schools are expected to take improvement action if they are shamed and blamed, for example (Elstad 2009). This is how evaluation systems are assumed to operate in an ideal world, but we do not know whether evaluation actually works in this way in local school governance.

One way to explore the significance of evaluation in this context is to examine how it is used in decision making. Unfortunately, discussion of evaluation use has often been confined by a "tacit normative framework" (Dahler-Larsen 2007, 20) favouring the intended and planned use of evaluation (Boswell 2009), which restricts our understanding of evaluation and evaluation systems in the real world. Other uses, functions and constitutive effects of evaluation merit equal attention.

In this article, *evaluation use* refers to individual actors' use of evaluation, while the *functions of evaluation* go beyond individual use and include, for example, legitimising school organisations or governance models. The *constitutive effects* of evaluation systems refer to how evaluation systems, for example, shape quality discourses and define what is important in education and school systems (Dahler-Larsen 2013; Hanberger 2013; Lingard and Sellar 2013; Power 1997).

We need better knowledge of evaluation in local school governance because evaluation-based decisions can have far-reaching consequences for schools, teachers and students. Schools producing poor results can be closed or given extra resources depending on decisions made by education committees, for example. A further argument is that there is dispute regarding the need for and value of various evaluation systems (Hanberger 2011; Leeuw and Furubo 2008), and whether they do more harm than good. Teachers are increasingly complaining that they are subjected

to too many accountability measures, which take time away from teaching and negatively affect education (Koretz 2009; Lingard and Sellar 2013). This article advances our understanding of how local governments respond to the globalisation of education governance in which evaluations are more and more important, and of how local governments cope with the accountability pressure and shape their own local evaluation systems and strategies.

The article considers the role of evaluation in local school governance at the central municipality level, seeking answers to three research questions: How are evaluations used and which functions do they have in local school governance? Do evaluations have constitutive effects and, if so, what are they? What explains how evaluations are used and the functions they have in this context?

As it is acknowledged that how evaluation systems are used depends partly on contextual factors, this article looks into what contextual factors matter in the studied context. If local school actors conceive standardised test scores as reflecting effectiveness and efficiency in education (Trujillo 2012) and if the culture of accepting and implementing state policies differs (Smith and Abbott 2014), responses to external evaluations can be expected to vary. Politicians from majority parties can be expected to use evaluations to justify and legitimise current policy in communities where education is highly polarised (Contandriopoulos, Lemire and Tremblay 2010), while opposition parties will use the same evaluations critically to expose failures.

Next, this paper develops the three key concepts used in the analysis. The cases explored, the methods and the material are then described. After that, we present, and then interpret, our findings concerning the use, functions and constitutive effects of evaluation in local school governance. The paper ends with conclusions and a discussion of the main uses, functions and constitutive effects of evaluation and of the importance of contextual factors for evaluation use.

## **2. Theoretical framework: the use, functions and constitutive effects of evaluation**

Three models of governance, i.e. the state, local government and multi-actor models, frame the analysis of the role of evaluation in *local school governance*. In this section, we develop and discuss the three key concepts and related theories used to explore the role of evaluation in local school governance.

The concept of *evaluation use* is multi-dimensional and refers, for example, to instrumental (e.g. decision support), conceptual (e.g. new insights) and symbolic (e.g. justification) uses. The literature reports that the relevance, quality and credibility of evaluations are three of the most important factors promoting evaluation use (Shulha and Cousins 1997). Cousins and Leithwood (1986, 348) also identified “communication quality”, “findings consistent with audience expectations”, “timeliness” (in disseminating results to decision makers), “political climate”, “personal characteristics” (cf. Askim 2008) and “commitment and/or receptiveness to evaluation”

as factors promoting evaluation use in policy settings. Learning from evaluation also depends on “engagement and interaction with new information and internal reflection on the meaning of that information” (Jacobson, Carter, Hockings and Kelman 2011, 55) and on sufficient time and opportunities for reflection and organisational support.

Evaluations can be used for *single- or double-loop learning* (Nilsson 2005). Single-loop learning is underpinned by the assumption that the key to improvement is conformity with predetermined goals. Double-loop learning is less certain about goals and assumes that they must be constantly re-examined, and modified or revised if needed. Single-loop learning implies learning to maintain the same standard of service (e.g. education) with less input or ensuring the same standard of service in all units, whereas double-loop learning implies changing the goals as needed to increase the capacity of an organisation to meet changing needs (Argyris and Schön 1996).

A survey study of Norwegian councillors’ use of performance information to inform political decisions (Askim 2008) indicates that *individual-level factors* explain differences in evaluation use, with “frontbencher”, “less-educated” and “inexperienced” councillors being the most inclined to seek and use performance information. Contandriopoulos *et al.* (2010), who systematically reviewed the relevant literature, defined *collective-level knowledge use* as how “users incorporate specific information into action proposals to influence others’ thought and practices” (459), finding three important contextual factors that influence this. First, in contexts with low “issue polarization” (similar to “political climate”), participants try to resolve problems using rational arguments whereas, in more polarised contexts, strategic participants “try to impose their views on others” (461). Second, the key actors (i.e. users, producers and intermediaries) “invest their energy and resources in knowledge exchange processes to the extent that they perceive this investment to be profitable” (462) and one or several of these actors must assume these costs in an attempt to reach a “cost-sharing equilibrium”. Users can invest resources in obtaining knowledge they find useful, producers devote resources to disseminating findings, while intermediaries (e.g. lobbyists) spend time advocating specific action proposals to defend or advance their interests (463). The third contextual factor, “social structuring”, acknowledges that “interpersonal trust facilitates and encourages communication and that repeated communication creates trust” (463) and that well-functioning formal or informal communication channels promote authentic knowledge exchange (468).

Further, school improvement can be promoted by an evaluation system if it is based on knowledge of factors facilitating school development. For example, evaluation systems designed to provide “assessment for learning” (William 2011) and applying positive sanctions are more likely to have learning and improvement effects (Elmore 2004). However, evaluation and evaluation systems can also have negative effects such as score inflation, teaching to the test, and emotional stress (Koretz 2009).

Important *functions* of evaluation in governance may not be captured by individuals' use of evaluation (Hanberger 2013). Besides instrumental functions such as accountability and school development, evaluations can perform two symbolic functions: legitimising organisations, and substantiating organisations' and actors' policy preferences (Boswell 2009, 7). Political organisations derive their legitimacy from “talk” and “decisions”, whereas action organisations derive theirs from output and performance (Boswell 2009, 47). Political organisations like education committees tend to use knowledge to support decisions and as a source of legitimacy, whereas action organisations such as school companies are more oriented to using evaluations to improve performance (Boswell 2009, 47) and earn profits. The framework pays attention to six functions that evaluation can have in local school governance: governance, school development, accountability, critical learning, legitimisation and symbolism (Hanberger, in this issue).

Evaluation systems may also have *constitutive effects*. Dahler-Larsen (2013, 8) recognises that performance indicators, the central components of “evaluation machines” (his term), have constitutive effects: “Effects of indicators are political as they define categories that are collectively significant in a society”. He maintains that three mechanisms shape what is considered important in a society: indicators produce order and reduce ambiguity in a messy world, provide a language and determine what should be measured and, in turn, determine what is measured constitutes what counts and is real, whereas what is unmeasured becomes invisible and unimportant. The constitutive effects of these mechanisms depend on contextual factors and stakeholder responses.

Similarly, Lingard and Sellar (2013) demonstrate the perverse and constitutive effects of high-stakes testing and of ‘catalyst data’ on teachers’ pedagogical practices. They claim that high-stakes testing data have systemically negative effects on schooling:

Performance data have catalyzed media and subsequent systemic reactions, but in the context of Australia’s federal system, such data are also contributing to a partial dissolution of the authority of State and Territory education systems, by constituting a national space of measurement, comparison and the associated allocation of funding (Lingard and Sellar 2013, 635).

Although the use, functions and constitutive effects of evaluation are treated separately in the above paragraphs, they are used in a complementary way to reflect and capture the role and significance of evaluation in local school governance.

### **3. Material and methods**

The studied municipalities were strategically selected in 2012 to ensure variation in contextual factors that may affect how external evaluations are managed and used and how local evaluation systems are set up (Table 1). The municipalities “East”, “West”

and “South”, with populations of 75,000–100,000 inhabitants, differ in their political majority (compliance with national government education policy and national evaluations can be expected when the political ideology is the same at both the municipal and national levels of government), school performance (low-scoring municipalities may be more inclined to use evaluation to take school improvement action) and share of independent schools (evaluation may be used more frequently to support informed school choice if the proportion of independent schools is high). “Share of students of non-Swedish ethnicity” and “parental education level” could also affect evaluation use. East scores higher than either West or South on all selection variables except for “share of students of non-Swedish ethnicity” where South stands out with a higher value. The last two background variables are often used to reflect conditions that affect school and student performance (Levin 2010) but may also hold implications for the use and functions of evaluation. We considered using other selection variables, such as “municipal size”, “municipal economy”, “criticism from Schools Inspectorate” and “type of local evaluation system”, assuming they might affect the use, functions and constitutive effects of evaluation. Municipalities of different sizes may set up or allocate resources for managing evaluations in various ways but, for exploring the present research questions, municipal size and economy seem relatively unimportant. Municipalities of the same size were selected as they deal with about the same number of schools. Unfortunately for our study, school inspections are not carried out during the same year in all municipalities and information about ongoing and recently completed inspections is unavailable; in addition, no national or local information is available about local evaluation systems. These two variables therefore had to be compiled by the researchers themselves.

The selection variables and the contextual variables discussed in the previous section are used in interpreting the results. In the analysis, we interpret the results in relation to all these variables but only discuss those that are clearly important in the material.

Table 1. Municipalities and selection variables (2012–2013 school year)

Selection variables	West	East	South
Political majority	Right–centre coalition with the Greens	Right-wing coalition	Coalition, mainly left-wing
% eligible for upper secondary school	78	95	76
Average grade, final year of compulsory school (national average 213.1)	203	242	204
Share of students in independent schools	20	26	19
Share of students of non-Swedish ethnicity	10	15	45
Share of pupils with at least one parent with post-upper-secondary education	57	73	44

Source: Information System on Results and Quality (SIRIS): [www.siris.skolverket.se/siris/](http://www.siris.skolverket.se/siris/) (Accessed 10 March 2014)

The article is based on municipal policy documents, minutes from municipal education committee meetings (2011–2013), school administration websites, and interviews with 15 well-informed people: three local government politicians from the majority party and two from the opposition party, seven administrators (i.e. Head of the Education Department, senior administrators, and evaluation experts), and three politically elected local auditors.

Interviewees were asked how and why they used evaluations and about the perceived main functions of evaluations in local school governance. The constitutive effects of evaluation were not cited in communication with the interviewees, but were applied in interpreting the interviews and documents. The interviewees were also asked how they responded to external evaluation systems, about their evaluation strategies, and how they valued and used data produced by national, international and local evaluation systems. Interviews, policy documents and minutes from education committee meetings were used to interpret how the web of evaluations was navigated.

Our approach is explorative and eclectic, borrowing elements from various perspectives. The analysis is guided by the three research questions. The minutes and interviews, which capture decision makers' experience and understanding of evaluation use, are referred to as first-level interpretations of evaluation use. Quotations exemplify issues at stake and how actors interpret the meaning and significance of evaluation. The researchers also interpret texts and interviews through the lens of theory as a second level of interpretation (Yanow 2000). In the discussion section, we elaborate on our second-level interpretations informed by evaluation use and (school) governance theory.

For simplicity we write, for example, "East is critical of the Schools Inspectorate", even though municipalities obviously do not constitute single actors. When variation clearly prevails within municipalities, this fact is highlighted. We do not refer to specific inspection reports to avoid revealing the identities of the municipalities.

## **4. Results**

### ***4.1 Characteristics of local evaluation systems***

Figure 1 summarises the web of evaluation systems identified in our material from the perspective of local decision makers. More or less all elements shown in Figure 1 constitute the three local evaluation systems identified here. The filter shows that decision makers selectively use and assimilate information from external evaluations. Locally developed elements are depicted on the left side of the figure.

East's evaluation system supports management by objectives and results-based management intended to enhance competition between schools, free school choice, and education quality. All elements shown in Figure 1 are in place and largely integrated into the local evaluation system. The system is set up to continually monitor the achievement of objectives and quality in education related to local strategic

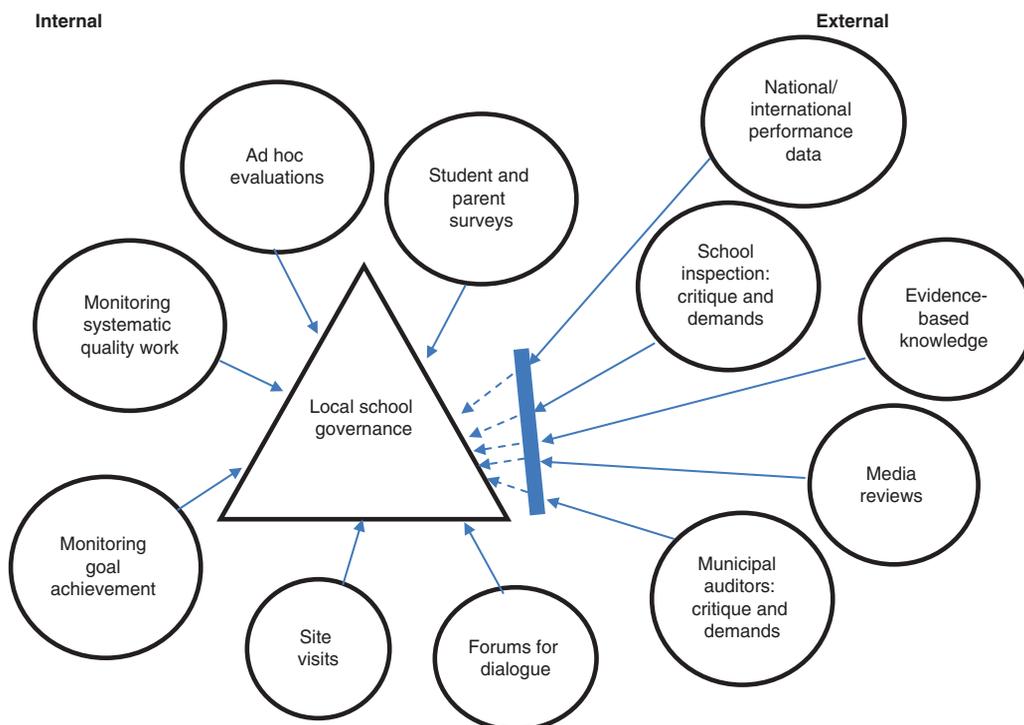


Figure 1. Local evaluation system: external evaluations are filtered and partly integrated into local school governance

objectives (i.e. maximal development, stimulating learning, real parent and student influence, and a safe working environment) and assessed against 18 national and local performance measures.

Evaluations are very important for the Education Committee to be able to continue to steer future development and gain insight. We have more than 100 pre-schools, 35 primary schools, and about 10 secondary schools, and with this number [of schools] it is impossible for politicians to be informed about what is going on in all schools, which is why monitoring and evaluation is something we really care about in the goal- and results-based governed East. (East/Interviewee 2)

A web-based benchmarking tool provides information about all schools, including about matters such as school ownership, student and parent satisfaction, and average grades. Parents and students can easily rank schools according to these variables. In addition, the website contains policy documents, minutes from Education Committees etc.

The monitoring and analysis of East's systematic quality work is the main pillar of the evaluation system. As a complement, East participates in inter-municipality evaluation cooperation and undertakes ad hoc evaluations of, for example, students not passing Grade 9, equity work, and school staff self-evaluations. Interviewees emphasise that ad hoc evaluations are important complements to the monitoring of systematic quality work.

International achievement tests, such as the Programme for International Student Assessment (PISA), are not discussed much, whereas performance data from Skolverket's (the National Agency for Education's – NAE's) SALSA tool and SIRIS system<sup>3</sup> are integrated into the local evaluation system.

The SSI has criticised East for insufficient monitoring and evaluation of its resource allocation system, while the municipal auditors have identified weaknesses in the governing and follow-up of East's own "school production" as well as cost overruns (East 2012). We will return later to East's responses to these evaluations.

In South, politicians and officials embrace the shared understanding that a quality management scheme constitutes the framework of local school governance and the evaluation system. While most elements shown in Figure 1 are apparently in place in the municipality, the interviewees describe the scheme as the lens through which the collection and use of evaluation data of various kinds is viewed.

A 2011 SSI report revealed that most South schools failed to meet national standards in several, if not most, key areas. For example, teacher qualifications, school-principal leadership skills, and teaching quality did not meet the standards, and a significant number of students, well above the national average, did not achieve basic educational goals. The SSI also found severe inadequacies in municipal school governance and in school leaders' monitoring and evaluation. Several of these failures were observed as early as 2005, some have even worsened since then, and an SSI inspection report concluded there was a "lack of well-defined education governance" in South. To improve the situation, the municipal auditors carried out an independent examination of the school principals' working conditions. A quality management scheme was then developed according to a common structure applying to all accountability holders. The scheme aims to: a) measure the performance and goal attainment of students and schools; b) analyse the current situation in light of several prioritised quality goals; and c) specify improvement measures. The quality goals are derived from national and municipal educational goals, and data about performance and goal attainment come primarily from student grades and national test results. Statistics about expenditures, as well as self-evaluations in which individual schools are accountable for goal fulfilment, are also used.

Today, the quality management scheme, regarded by the interviewed politicians and administrators as the most important pillar of the evaluation system, frames recurrent dialogues between the Education Department and school principals. Information about aspects measured by the scheme is reported annually to the Education Committee, municipal assembly, and executive committee. According to the interviewees, every decision affecting schools requires compulsory feedback:

The quality management scheme has resulted in an extensive feedback system for results, but also for planned and implemented improvement measures. In this way, the Committee gets a good grasp of the 'numerical' results, but perhaps not such a good grasp of things that can't be measured. (South/Interviewee 2)

To support school choice, South provides web-based information similar to that provided by East, including about the student–teacher ratio, expected performance, and parent and student satisfaction with municipal schools (but not independent schools). It also provides links to the SIRIS system, access to the systematic quality work scheme, minutes from Education Committee meetings, and other policy documents.

South strives to implement evidence-based practices by informing managerial decisions and organisational practices with the best available scientific evidence concerning, for example, factors for successful schools and education leadership. Teachers are also encouraged to conduct small-scale pilot experiments to evaluate the effects of certain types of teaching in order to improve students' school results.

Similarly, West's evaluation system supports management by objectives and all elements in Figure 1 are more or less in place. The local evaluation system is based on an annual evaluation cycle that involves the municipal, school and student/parent levels. This cycle is an important component of the municipality's education governance, budgeting and development work. The quality seminars, at which politicians, school principals and key administrators meet twice a year, are central fora in West's local school governance. The seminar dialogues focus on the goal achievement specified in the goal and resource plan, which encompasses the national educational goals and 35 indicators concerning comprehensive schooling. External evaluations are filtered before being considered in local school governance:

Management by objectives is a good model because if you focus on the ultimate goal, you cannot be much of a 'flip-flopper'. Even if there is a new PISA report or the NAE publishes new reports, we cannot change our goals, but they [i.e. the reports] may provide input for our analysis. (West/Interviewee 2)

However, West is dissatisfied with its evaluation system:

We lack tools for evaluation analysis, which is critical for an evaluation worthy of the name. Right now, we follow up and report but we lack the rest of the cycle needed so that our new goals and resource distribution can function in the way we want. (West/Interviewee 2)

In addition, various ad hoc evaluations are conducted to complement the quantitative evaluations. For example, an evaluation examining school principals' working conditions has helped relieve the school principals of certain tasks, and an evaluation aimed at developing "success factors" resulted in "future-directed changes, not the same old ones but a new organisation and success factors", according to the Head of the Education Department.

The school administration works intensively to address deficiencies in the school sector, partly in response to heavy criticism from the Schools Inspectorate in 2011.

The three evaluation systems differ somewhat, but all provide indispensable information for systematic quality work and local school governance. The local

decision makers describe their evaluation systems as aligned with the ideal of outcome-oriented school governance but, as demonstrated below, this is not the only way to perceive these evaluation systems. The filtering of external evaluations (Figure 1) indicates that local governments use their discretion in the governance structure and that compliance with national evaluation demands and governing does not occur without judging the relevance of the information to local school governance.

## **4.2 Use and functions of evaluation**

### *Decision makers' uses of evaluations*

Evaluations are used in many ways, although there are some common patterns in how the political majority and the administration use evaluations in the three studied municipalities. Evaluations are employed to identify problems related to poor performance and to support decisions aimed at school improvement. They are also used to follow and analyse trends in school/student performance, parent/student satisfaction with schools, and staff satisfaction with working conditions:

I would say that all of our evaluations are very important in governing school improvement, and for that we need a really good basis . . . Our quality reports, evaluations and in-depth analysis are the basis for our school governance. (East/Interviewee 1)

A good evaluation identifies success factors that we can refine, strengthen and distribute at individual, group, unit, department and municipal levels. (West/Interviewee 1)

Further, some key performance measures are used to benchmark schools and the municipality.

We use our evaluations to find out how we perform and [to determine] our position, and we use them strategically to identify our strengths and areas in need of improvement. (East/Interviewee 2)

Decision makers use evaluations when holding dialogues with school leaders and when informing parents and students to facilitate free school choice.

I look at evaluations – mainly customer surveys, in-depth evaluations, and Schools Inspectorate reports – read them carefully with my colleagues, and we analyse the results, trends, differences, look for reasons, and consider what needs to be changed, what action to take – these types of questions. (East/Interviewee 3)

It is important that we base our quality seminars on analysis; otherwise, they just become talk clubs. (West/Interviewee 2)

Evaluations are also used for accountability, mainly to hold school principals and schools to account for school performance, quality work, and budgeting. Evaluation

can also be used to confirm and justify school policy and to promote the municipality – also observed in our material – reflecting a political logic that may undermine the validity of the evaluation knowledge used.

The interviewees demonstrate that evaluations have many overlapping functions in local school governance. The decision makers give different weights to various components of the evaluation system (Figure 1), but all of them emphasise that many components must be used together for monitoring, analysing trends, and benchmarking.

The state continuously inspects and holds local governments accountable for school performance (Hanberger, in this issue). Before a school inspection report was used, decision makers analysed it carefully, and responded to it or complied. East's decision makers, for example, responded strongly to the SSI report mentioned earlier. They questioned how the inspection was conducted as well as whether it accurately reflected the local school governance and the actions taken to alleviate problems in East. The response was developed in a letter formally responding to the Head of the Schools Inspectorate.

... we were not at all impressed with how the inspectors operated, so I ended up writing a letter to the Head of the Inspectorate about what they were doing, saying that we found it difficult to understand how they interpreted their commission and that we did not understand what they wanted from us. They conducted strange interviews that were difficult to respond to ... we are not always sure that the SSI is right and we have to learn how to change our relationship with the SSI ... we need to balance our criticism and foster good relations with the SSI. This was done through the letter in which we said that they needed to improve to be taken seriously – sending us a report that does not say much is not worth spending time on. (East/Interviewee 2)

The inspection also upset the Chair of the Education Committee. The Committee responded to the SSI mainly by justifying how they had striven to reduce segregation in East, indicating that the SSI had not understood this. Similarly, the response to the critical audit report mentioned above was to downplay its significance. An earlier audit report criticising the local school governance model and unclear division of power and responsibility was also responded to in a similar way. Moreover, the decision makers' response to local auditors who repeatedly visited the Education Committee was to listen to the auditors' criticisms, learn about their mandate, and consider whether or not action was needed in response to the criticisms.

The use of evaluations in South is organised according to a quality management scheme that establishes the kinds of evaluative information to be collected, analysed and communicated. Statistical information about student grades, national test results, school self-evaluations and SSI reports are the main sources of information in the scheme, which is also formally integrated into decision-making processes at all levels. The development and use of the scheme is itself a consequence of a 2011 SSI report.

When evidence-based knowledge informs policy and practice in South, for example, in the case of managerial decisions and organisational practices, it is a sign of the use of evaluation.

Administrators in South do not consider international tests (e.g. PISA) or national ranking systems (e.g. SALAR's<sup>4</sup> Open Comparisons and those produced by the Swedish Teachers' Union) useful for local school governance and education improvement. They admit, however, that politicians are highly sensitive to both international test results and national rankings, especially when the media reacts strongly.

There are also concerns over the validity, value and usefulness of South's user-satisfaction surveys distributed annually to all students, parents and school staff. Such surveys cannot say whether individual students, parents and staff are better or worse off at any given time, and cannot be used as a basis for decision making.

What we need is very different from those superficial user-satisfaction surveys. We would like to know more down-to-earth things, such as whether and how certain kinds of lessons make any difference to the students. What did you learn in school today? Have you learned or understood something that you didn't before? This kind of more qualitative feedback information could be used directly by teachers to make school activities better. (South/Interviewee 1)

West emphasises using evaluations to support school governance, budgeting and development. The goal and resource plan, which is based on evaluation, systematic work and research, is used as the main governance tool. The Chair of the Education Committee says that the primary intention of evaluation "is about improvement". According to its Head, the Education Department strives to "conduct smart analyses" that can serve as a basis for ongoing school development. He regards school principals as key actors in realising educational goals: "The principal is crucial for the realisation of the school's task". He also emphasises governing: "We want to establish systematic governing and management of the school's tasks at all levels, that is, to educate our principals to create goals, structures, processes and analyses" (West/Interviewee 1).

External evaluations and international measurements do not play an important role in West, although SALAR's Open Comparisons is an important information source in the systematic quality work, not least in benchmarking. The Education Department Head says that the Open Comparison system "is objective and factual and facilitates comparisons" (West/Interviewee 1). Further, although the municipality did not fully agree with the SSI's critique, it nevertheless influenced local education policy and was cited when the requested changes were being made.

### *The opposition's use of evaluation*

The interviewee with the opposition in East questions whether the consequences of East's governance model for equity/segregation have been evaluated, and whether

the evaluation system produces knowledge that helps in understanding the causes of variable school performance. The opposition has also a different view of how those in power use evaluations:

If you get a critical evaluation they say, like, “Yes, but we still perform better than other municipalities and we are not worse than others”. But if you get an evaluation that shows that East is good, then they make a big deal about this . . . yes, confirmation that you are doing the right thing, but critical evaluations are not paid much attention. (East/Interviewee 4)

Moreover, the interviewee also questions the Education Committee’s formal response to the SSI. In expressing their reservations, the opposition said that they

preferred a response based on the SSI’s recommendations instead of on the speech defending the current [resource allocation] system that the Board had decided on. Therefore, we decided not to participate in the decision, neither this time, nor when the system was instituted. (East 2013)

Representatives of the opposition parties in South rejected the invitation to be interviewed. However, close examination of the minutes of Education Committee meetings over the previous two years indicates that, although the opposition parties had occasionally criticised local school governance and evaluation, the political context seems less polarised than that of East or West (see below). In South, evaluations were not used as political ammunition (Weiss and Bucuvalas 1977).

The opposition politician in West questions the validity of the current evaluations and emphasises the importance of connecting the evaluations to the teachers. He does not think that Open Comparisons’ numbers are sufficient, for example, and claims that statistics are often simply used for legitimation. To gain a realistic understanding of schools, he uses informal contacts to obtain information and penetrate the smokescreen that school principals sometimes attempt to create. Informal channels “are the way we gather information and take the pulse of everything”. He also thinks that the Swedish Teachers’ Union rankings are useful. The opposition politician is concerned because he thinks the authorities have little insight into the results and operations of independent schools, even though they occasionally visit them: “We do not have a grasp of that at all . . . we have no evidence about the work of these schools” (West/Interviewee 3). He thinks that the Education Committee does not have a transparency strategy and, since the municipal school budget is public, he would like the finances of the independent schools to be public as well.

The interviewed opposition politicians from East and, to some extent, from West are critical of how the local evaluation system is used by the political majority, indicating they are more polarised school communities than that of South.

### *Main functions of evaluation*

The three municipalities' evaluation systems support the five main functions of local school governance; namely, accountability, governance, school development, single-loop learning, and legitimation.

For the political majority and the administration, the local evaluation system in East serves chiefly to sustain and legitimise the governance model, support governance, and hold schools accountable for goal achievement and quality work:

The web platform allows citizens to see all the evaluations and allows parents who need special support for their children to get information and, together with our analysis and conclusions, helps create greater acceptance of our school policy and governance model. Credibility and accessibility are important even though not all parents use the information to make informed school choices. (East/Interviewee 1)

In South, the local evaluation system revolves around the newly established quality management scheme (see above) whose primary function is to support governance and goal achievement and to hold schools somewhat accountable. However, considering the relative newness of the scheme and the serious and long-lasting problems identified by the SSI in 2011, it is doubtful whether the scheme has taken root and clearly fulfils those functions in South.

The main function of evaluations in West is to inform the budget, support governance and school development, and facilitate goal achievement. The national evaluation systems often have a symbolic function: "It is often just about filling in forms and then it is forgotten" (West/Interviewee 3).

### **4.3 Constitutive effects**

Considering their more tacit and indirect effects, the local evaluation systems also reinforce a notion of parents and students as customers and the concept of a school market, and that customers should make repeated school choices to ensure the best possible education. The monitoring of key performance measures has a related constitutive effect in that it inculcates the norm that benchmarking is natural and worthwhile. Customers rank schools and schools compete and benchmark themselves: "I feel there is a need and call to compare oneself with others and over time in the organisation" (East/Interviewee 6).

The national evaluation systems, in particular the SSI, reinforce the local decision makers' role as the ultimate accountability holders in the national education system. The local evaluation system reinforces the school principals' and schools' role in the education system as street-level accountability holders relative to all other accountability holders (i.e. the SSI, education committees, local school administrations, local auditors, and parents and students). Hence, both national and local evaluation systems have constitutive effects in terms of establishing and reinforcing who is accountable for school and student performance.

The use of key performance indicators and the focus on measuring grades and results in national tests have constitutive effects in that these performance measures shape what is deemed worth knowing about school performance and define what constitutes quality in education (Dahler-Larsen 2012a, 2013). In the studied municipalities, the percentage of students passing Grade 9 and national test results are conceived as reflecting high performance and quality in education. Most interviewees are conscious of some problematic aspects of evaluation (e.g. focusing on what is measurable) and that multiple factors explain school and student performance but, when it comes to what is actually considered quality, grades and test results are crucial. For example, the Head of the Education Department in West states that the average final grade of compulsory school “is the most important output to highlight for the school provider, the environment and ourselves”. This demonstrates the constitutive effect of a key performance measure and the risk of reductionism (Dahler-Larsen 2012b), that is, evaluation of the curriculum being narrowed to measurable goals, while broad qualitative goals disappear.

Although PISA results are not used in local school governance, the PISA programme has a similar constitutive effect in that it defines quality in education and education systems, and the skills and competencies citizens need in a globally competitive world (Hanberger 2014).

The quite rigid character of the quality work scheme in South has led to a pragmatic, sceptical view of evaluation information, reinforced by the over-proliferation of evaluations. This implies that a new norm has evolved, i.e. scepticism regarding the validity of evaluations before using them.

The strong audit culture fostered by the governance and evaluation model has also, according to one opposition politician, fostered a culture of silence, particularly in the independent schools where “teachers are not allowed to criticise their own schools” (West/Interviewee 3). This is an example of a negative constitutive effect.

Another effect of monitoring and evaluating performance identified in the interviews is that decision makers’ focus on what does not work yields negativism (Schillemans and Bovens 2011). The fixation on poor results and problems creates frustration among teachers in problem schools.

Some schools feel frustration when we ask questions about negative things, and wonder why we do not ask about what works and what is positive. We try to do both, but a single school will not necessarily get questions about both negative and positive things. (East/Interviewee 1)

## **5. Conclusions and discussion**

Three main conclusions can be drawn from this study. First, local decision makers develop the evaluation capacity they need to manage the web of evaluations, formulate the response strategies they require to manage external evaluations, and set up their own evaluation systems to support local school governance.

Second, the use of evaluations in local school governance differs between the majority and opposition parties, between municipalities with high and low school performance, and between municipalities inclined or not inclined to support NPM and school choice.

Third, some key performance indicators from the National Agency of Education and SSI inspections affect local school governance in that they define what is considered important in education and reinforce the norm that benchmarking is natural and worthwhile, indicating the constitutive effects of national evaluation systems.

Local evaluation systems are set up to sustain local school governance and ensure compliance with the Education Act and other state demands. Local decision makers have learned to navigate the web of evaluations and developed response strategies for dealing with external evaluations so they can take into account what is useful and what cannot be ignored in order to avoid sanctions. Evaluation strategies are developed in response to the growing accountability pressure, particularly that from the SSI. East has the most explicit and developed evaluation strategy, which includes use of national data considered valuable, complemented with data collected through their own evaluations. Generally, the most trustworthy and authentic information about local schools, however, is the information decision makers obtain first-hand from site visits or through informal channels. This implies that knowledge produced by evaluation systems needs further validation and legitimation to be accepted as revealing truth about schools and schooling. The threat of getting entangled in the web of evaluation systems is effectively managed by decision makers, although it may still be a real problem for school leaders and teachers, for example.

Evaluations serve many functions in local school governance, the main ones identified here being maintaining and legitimising the applied governance model and supporting governance, accountability and school development in the interest of goal achievement. Certain national key performance indicators (e.g. student grades and results on national tests) inform local school governance and are taken at face value, being integrated into the goal structure and used for benchmarking. This conclusion confirms the recognition that the NAE's SIRIS system and SALSA tool contain important performance measures (Hansen and Lander 2009) that define what is important in education, suggesting a constitutive effect (Dahler-Larsen 2013) and that governing by numbers (Lingard 2011) is increasing. How the school community performs also affects how external evaluations are used and which functions they are allowed to serve in local school governance. Perverse effects of catalyst data<sup>5</sup> (Sellar and Lingard 2013), e.g. when used to "... improve or maintain the reputation of schools and systems and to secure funding, rather than the intended objective of improving literacy and numeracy outcomes in schools", do not prevail in our material. Improving school results is more conceived as an effective way of improving and maintaining reputation. Although we have demonstrated ways

to manage the accountability pressures, SALSA and SIRIS and the SSI have also contributed to strengthening the authority of the State.

All three municipalities have institutionalised fora for exchanging evaluation knowledge in their line organisations. This kind of social structuring (Contandriopoulos *et al.* 2010) promotes the use of evaluation for accountability, single-loop learning, and related governance and school development purposes, focusing on improving student performance and education quality according to pre-set goals or targets. The knowledge exchange and dialogues are confined to those invited to participate (not the opposition). Local evaluation systems are thus used mainly as management support systems (Dahler-Larsen 2000) and to confirm rather than question policies (Leeuw and Furubo 2008), but when evaluations are discussed in education committees the political dimension of evaluations become visible (Cousins and Leithwood 1986; Hanberger 2011; Weiss and Bucuvalas 1977).

The study supports the finding that ‘issue polarisation’ (Contandriopoulos *et al.* 2010) is an important contextual factor affecting evaluation use, as reported in the literature. How the municipal political majority and opposition use, respond to and question evaluative knowledge produced by national and local evaluation systems indicates a polarised context. The political majorities in East and West, which both have right-wing local governments, have great confidence in the new public management models of school governance, while the opposition has little confidence in them. This polarisation played out when East’s Education Committee questioned the validity of the SSI report, while the opposition agreed with the report’s recommendations. Our results confirm that, in contexts with high issue polarisation (Contandriopoulos *et al.* 2010), the use of evaluation differs between the political majority and opposition and that responses to external evaluations follow the same pattern. These differences in evaluation use do not primarily reflect party differences, but reflect who is in power and the knowledge needs of the applied governance model. Management by objectives and results shapes the use of evaluations to facilitate governance, accountability and school development targeting national and local objectives.

The SSI has criticised all three municipalities, each of which responded differently: East questioned the relevance of the inspection, whereas South and West (with some reservations) accepted the critiques and complied with them. This indicates that the accountability function works differently in the three municipalities, possibly due to the contextual factors *confidence in the governance model and evaluation system* and *school performance* and to the *self-perceptions of the municipalities as school communities*. East appeared to have the greatest confidence in its governance model and evaluation system which, together with its high performance in national rankings, gave it the confidence to criticise the SSI and to exploit positive evaluations for promotional purposes. In contrast, the SSI’s critiques went unchallenged in South and West and positive evaluations were not used for promotion. This interpretation is in line with Keddie’s (2013) recognition of how high-performing secondary schools in

the London area “. . . do thrive amid the external contingencies of the audit culture – they are able to fashion a triumphant and outstanding identity” (764). High performance confers status and builds confidence sufficient to deal with external demands and navigate the evaluation web at various levels of school governance.

The results indicate that a municipality with high school performance responded critically to the SSI (East), whereas a municipality repeatedly criticised by the SSI was more inclined to comply without objection (South). The confidence engendered by high school performance also results in support for the use of evaluations to promote and legitimise current school policy and for the applied governance model. The assumption that the “share of students in independent schools” affects evaluation use finds some support in that East’s local evaluation system is set up to facilitate and support informed school choice more than the systems of West and South are.

One’s view of the governance doctrine under scrutiny and one’s vision of evaluation in democratic governance can influence one’s interpretation of the functions and constitutive effects of evaluation in local school governance. For example, advocates of deliberative democracy would deem an evaluation system that reinforces the role of citizens (i.e. students and parents) as customers in contrast to promoting active citizenship as a negative constitutive effect.

The article improves our knowledge of how local decision makers build their evaluation capacity, of how and why they use evaluations in practice, and of the constitutive effects of evaluation systems.

We hope this analysis of the role of evaluation in local school governance in Swedish municipalities advances research into this issue and that it can inspire more in-depth research into how decision makers navigate the web of evaluation systems, how they reason about the quality and relevance of evaluation information, and how contextual factors affect the use and functions of such information. Not all the selection variables and contextual variables identified here are discussed in detail because they were not all clearly evident in the material; however, these and other contextual variables still merit further attention in future research as they could affect the role of evaluation in local school governance. There is also a need to explore the constitutive effects of evaluation because such effects are paid little attention in the policy discourse. Evaluation systems can redefine what is important in education, reshape schools, teaching, identities and the lives of school principals, teachers and students in ways that merit much greater examination and discussion.

---

*Anders Hanberger* is Professor in Evaluation and director of research at the Umea Centre for Evaluation Research, Umea University. His research areas cover policy analysis and evaluation research and focus on public policy, governance, evaluation, evaluation systems and methodology. He has a special interest in the role of evaluation in democratic governance.

*Lena Lindgren* is Associate Professor at the School of Public Administration, Gothenburg University. She works with teaching and research, focusing on evaluation and policy analysis, and has also conducted several evaluations on commission for government agencies, particularly in the field of education.

*Ulf Lundström* is Associate Professor at the Department of Applied Educational Science, Umeå University. His main research interests lie within the areas of education policy, e.g. focusing on marketisation and inclusion, evaluation and the teaching profession.

## Notes

- 1 *Evaluation system* refers to “the procedural, institutional and policy arrangements shaping the evaluation function and its relationship to its internal and external environment” (Liverani and Lundgren 2007, 241). Evaluation system also refers to routines established for dealing with stand-alone evaluations and to a system producing streams of evaluation information.
- 2 *Local school governance* refers to governance that occurs in a municipality and in a quasi-market where local school actors govern and influence schooling and education. It includes the efforts of actors and institutions to govern and influence matters such as school policy, education, school climate and school safety.
- 3 SIRIS stands for Skolverkets Internetbaserade Resultat- och kvalitets Informations System (the NAE’s Internet-based results and quality information system) while SALSAs stands for Skolverkets Arbetsverktyg för Lokala SambandsAnalyser (the NAE’s tool for local correlation analysis).
- 4 SALAR stands for the Swedish Association of Local Authorities and Regions.
- 5 Performance measures from a national test (National Assessment Programme – Literacy and Numeracy, or NAPLAN) are used as a performance measure for state education systems (Sellar and Lingard 2013).

## References

- Act 2010:800. *Skollag* [Education act].
- Andersen, V. N., Dahler-Larsen, P. & Strömbæk Pedersen, C. (2009). Quality assurance and evaluation in Denmark. *Journal of Education Policy*, 24(2), 135–147.
- Argyris, C. & Schön, D. (1996). *Organizational learning II: theory, method and practice*. Reading, MA: Addison-Wesley.
- Askim, J. (2008). The demand side of performance measurement: explaining councillors' utilization of performance information in policymaking. *International Public Management Journal*, 12(1), 24–47.
- Boswell, C. (2009). *The political uses of expert knowledge: immigration policy and social research*. Cambridge, UK: Cambridge University Press.
- Contandriopoulos, D., Lemire, M., Denis, J.-L. & Tremblay, É. (2010). Knowledge exchange processes in organizations and policy arenas: a narrative systematic review of the literature. *Milbank Quarterly*, 88(4), 444–483.
- Cousins, J. B. & Leithwood, K. A. (1986). Current empirical research in evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
- Dahler-Larsen, P. (2000). Surviving the routinization of evaluation: the administrative use of evaluations in Danish municipalities. *Administration and Society*, 32(1), 70–92.
- Dahler-Larsen, P. (2007). Constitutive effects of performance indicator systems. In *Dilemmas of Engagement: Evaluation and the New Public Management*, S. Kushner and N. Norris (eds.), 17–35. Amsterdam, Elsevier.
- Dahler-Larsen, P. (2012a). *The evaluation society*. Stanford, CA: Stanford Business Books.
- Dahler-Larsen, P. (2012b). Evaluation as a situational or universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look in practice, and why it is fashionable. *Scandinavian Journal of Public Administration*, 16(3), 29–46.
- Dahler-Larsen, P. (2013). Constitutive effects of performance indicators: getting beyond unintended consequences. *Public Management Review*, 16(7), 969–986.
- de Wolf, I. F. & Janssens, F. J. G. (2007). Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*, 33(3), 379–396.
- East. (2012). Economic and organizational steering in municipal school production. *Audit report 3/2012*. East municipality: Sweden.
- East. (2013). *Minutes of Education Committee meeting, 16 May 2013*. East municipality: Sweden.
- Elstad, E. (2009). Schools which are named, blamed and shamed by the media: school accountability in Norway. *Educational Assessment, Evaluation and Accountability*, 21(2), 173–189.
- Ehren, M. C. M., Altrichter, H., McNamara, G. & O'Hara, J. O. (2013). Impact of school inspections on improvement of schools: describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability*, 25(1), 3–43.
- Elmore, R. F. (2004). Conclusion: the problem of stakes in performance-based accountability systems. In *Redesigning Accountability Systems for Education*, S. H. Fuhrman and R. F. Elmore (eds.), 274–296. New York: Teachers College Press.
- Hanberger, A. (2011). The real functions of evaluation and response systems. *Evaluation*, 17(4), 327–349.
- Hanberger, A. (2013). Framework for exploring the interplay of governance and evaluation. *Scandinavian Journal of Public Administration*, 16(3), 9–28.
- Hanberger, A. (2014). What PISA intends to and can possibly achieve: a critical programme theory analysis. *European Educational Research Journal*, 13(2), 167–180.

- Hanberger, A. (2016). Evaluation in local school governance: a framework of analysis. *Education Inquiry*, in this issue.
- Hansen, M. & Lander, R. (2009). Om statens verktyg för skoljämförelser. Vem vill dansa SALSA? [The state's tools for school comparisons. Who wants to dance the Salsa?]. *Pedagogisk forskning i Sverige*, 14(1), 1–21.
- Jacobson, C., Carter, R. W., Hockings, M. & Kelman, J. (2011). Maximizing conservation evaluation utilization. *Evaluation*, 17(1), 53–71.
- Keddie, A. (2013). Thriving amid the performative demands of the contemporary audit culture: a matter of school context. *Journal of Education Policy*, 28(6), 750–766.
- Koretz, D. (2009). *Measuring up: what educational testing really tells us*. Cambridge, MA: Harvard University Press.
- Leeuw, F. (2003). Reconstructing program theories: methods available and problems to be solved. *American Journal of Evaluation*, 24(1), 5–20.
- Leeuw, F. & Furubo, J.-E. (2008). Evaluation systems: what are they and why study them? *Evaluation*, 14(2), 157–169.
- Levin, B. (2010). Governments and education reform: some lessons from the last 50 years. *Journal of Education Policy*, 25(6), 739–747.
- Lingard, B. & Sellar, S. (2013). “Catalyst data”: perverse systemic effects of audit and accountability in Australian schooling. *Journal of Education Policy*, 28(5), 634–656.
- Lindgren, L., Hanberger, A. and Lundström, U. (2016). Evaluation systems in a crowded policy space: implications for local school governance. *Education Inquiry*, this issue.
- Liverani, A. & Lundgren, H. (2007). Evaluation systems in development aid agencies: an analysis of DAC peer reviews, 1996–2004. *Evaluation*, 13(2), 241–256.
- Merki, K. M. (2011). Special issue: accountability systems and their effects on school processes and student learning. *Studies in Educational Evaluation*, 37(4), 177–179.
- Mintrop, H. & Trujillo, T. (2007). The practical relevance of accountability systems for school improvement: a descriptive analysis of California schools. *Educational Evaluation and Policy Analysis*, 29(4), 319–352.
- Nilsson, M. (2005). The role of assessments and institutions for policy learning: a study on Swedish climate and nuclear policy formation. *Policy Sciences*, 38(4), 225–249.
- OECD. (2011). *OECD reviews of evaluation and assessment in education, SWEDEN*. Paris: OECD Publishing.
- Power, M. (1997). *The audit society: rituals of verification*. Oxford, UK: Oxford University Press.
- Ravitch, D. (2010). *The death and life of the great American school system: how testing and choice are undermining education*. New York: Basic Books.
- Schillemans, T. & Bovens, M. (2011). The challenge of multiple accountability: does redundancy lead to overload? In *Accountable Governance: Problems and Promises*, M. J. Dubnick and H. G. Frederickson (eds.), 3–21. Armonk: ME Sharpe.
- Shulha, L. & Cousins, J. B. (1997). Evaluation use: theory, research and practice since 1986. *Evaluation Practice*, 18(3), 195–207.
- Smith, P. & Abbott, I. (2014). Local responses to national policy: the contrasting experiences of two Midlands cities to the Academies Act 2010. *Educational Management Administration and Leadership*, 42(3), 341–354.
- Trujillo Tina, M. (2012). The disproportionate erosion of local control: urban school boards, high-stakes accountability, and democracy. *Educational Policy*, 27(2), 334–359.
- Weiss, C. H. & Bucuvalas, M. (1977). The challenge of social research to decision making. In *Using Social Research in Public Policy Making*, C. H. Weiss (ed.), 213–234. Lexington, MA: Lexington Books.

- Whitby, K. (2010). *School inspection: recent experiences in high performing education –systems - literature review*. Reading, UK: CfBT Education Trust.
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14.
- Yanow, D. (2000). *Conducting interpretive policy analysis*. Newbury Park, CA: Sage.